A Hybrid Machine Learning Model for Predicting Surgical Procedure Duration: Integrating Random Forest and K-Means Clustering

Amira Brahmi [1], Asma Ouled Bedhief [2], Safa Bhar Layeb [3], Najla Aissaoui [4]

Abstract— **Efficient operating room (OR) management depends on the accurate prediction of surgical procedure durations to improve scheduling, enhance patient outcomes, and reduce operational costs. This study presents a hybrid machine learning model that combines Random Forest and K-Means clustering to predict the duration of cholecystectomy procedures. The model is trained using real-world data from the digestive surgery department at Mahmoud El Matri Hospital in Tunis, Tunisia, incorporating patient demographics, surgeon experience, and other contextual factors. Synthetic data generation was also applied to reinforce model reliability. The proposed approach achieved strong performance, with a root mean square error (RMSE) of 0.45 minutes, a mean absolute error (MAE) of 0.36 minutes, and a coefficient of determination (R^2) of 0.99. Comparative analysis with individual models such as Random Forest, K-Means, decision trees, and linear regression confirms the hybrid model's superior predictive capability. These results demonstrate the potential of the proposed hybrid model as a practical tool for optimizing OR scheduling and improving healthcare resource management.**

I. Introduction

Operating rooms (ORs) are the most critical and resource-intensive units in any hospital, representing both the largest source of revenue and the most significant cost center [1]-[2]. As such, optimizing OR operations is a high-priority challenge for healthcare systems worldwide. In addition to the financial implications, improving OR efficiency can lead to substantial benefits, including reduced patient waiting times, improved staff satisfaction, enhanced resource utilization, and better clinical outcomes [3].

Central to OR efficiency is the accurate prediction of surgical case durations, which directly affects scheduling precision, operating room turnover, and post-operative care coordination. Traditional estimation methods based on historical averages or surgeon experience often fail to account for the complexity and variability inherent in surgical procedures. This has spurred increasing interest in the application of machine learning (ML) techniques to support decision-making in surgical scheduling.

In the era of data-driven healthcare, ML has emerged as a powerful tool to address various OR challenges—ranging from predicting surgery durations and optimizing schedules to anticipating cancellations and managing recovery units [4]-[5]. Algorithms such as Random Forest, XGBoost, and neural networks have demonstrated notable success in this context, offering more accurate and dynamic predictions than traditional approaches [5]. ML models have also been employed to coordinate the use of interconnected care units, such as the post-anesthesia care unit (PACU) and ORs, enabling more integrated and efficient hospital operations [6]. Among these algorithms, Random Forest stands out for its robustness, tolerance to outliers, and resistance to overfitting, making it highly suitable for clinical prediction tasks involving complex and noisy data [2], [7]. It has been effectively used to predict surgery durations, identify high-risk cancellations, and forecast emergency patient arrivals [1]; [8]. Comparative studies also confirm the superiority of Random Forest over other models like linear regression and multivariate adaptive regression splines, especially in terms of accuracy and generalizability [9].

However, many existing models rely on a single machine learning algorithm, which can lead to limitations such as poor generalization, reduced interpretability, and high computational complexity. To address these shortcomings, this study proposes a hybrid machine learning model that integrates supervised and unsupervised learning techniques, building on the approach of Amrollahibiouki and Beauregard [10], who developed a hybrid model combining Random Forest and K-means clustering to predict surgical durations at the McGill University Health Centre, achieving an average prediction error of 4.6 minutes and surpassing the performance of standalone models. This combined strategy enhances prediction accuracy by leveraging clustering to group similar surgical cases, thereby improving the model's robustness and effectiveness.

The proposed hybrid model combines Random Forest regression with K-Means clustering to predict the duration of cholecystectomy procedures. Developed using real-world data from the digestive surgery department at Mahmoud El Matri Hospital in Tunis, Tunisia, the model incorporates patient-specific factors, surgeon experience, and synthetic data to enhance its reliability. Through rigorous testing, we

1: LR-OASIS, Department of Industrial Engineering, National Engineering School of Tunis, University of Tunis El Manar, Tunis, Tunisia
Email: amira.brahmi@enit.utm.tn
2: LR-OASIS, Department of Industrial Engineering, National Engineering School of Tunis, University of Tunis El Manar, Tunis, Tunisia
Department of Industrial Engineering, National Engineering School of Bizerte, University of Carthage, Tunis, Tunisia
Email: asma.ouledbedhief@enit.utm.tn
3: LR-OASIS, Department of Industrial Engineering, National Engineering School of Tunis, University of Tunis El Manar, Tunis, Tunisia
Centre Génie Industriel, Université Toulouse, IMT Mines Albi, Albi, France
Email: safa.layeb@enit.utm.tn
4: LR-OASIS, Department of Industrial Engineering, National Engineering School of Tunis, University of Tunis El Manar, Tunis, Tunisia
Department of Industrial Engineering, National Engineering School of Carthage, University of Carthage, Tunis, Tunisia
Email: najla.aissaoui@enicar.ucar.tn

aim to demonstrate its superior performance compared to standalone models, offering a practical tool to improve OR scheduling and enhance healthcare delivery. The remainder of the paper is structured as follows: Section II presents the methodology, Section III describes the principles of the predictive model, Section IV discusses the results, and the final section concludes the paper and suggests directions for future research.

II. METHODOLOGY

A. Data Collection

Data were gathered from the digestive surgery department at Mahmoud El Matri Hospital in Tunis, Tunisia, with a focus on cholecystectomy procedures, the most frequently performed surgery during the study period. Initial data collection faced challenges, including inconsistent procedure scheduling (e.g., no cholecystectomies on certain Wednesdays) and restricted access to electronic medical records.

To overcome these limitations and ensure a robust dataset, synthetic data were generated based on observed procedure durations. Using the Input Analyzer in ARENA software, real data were fitted to an initial Beta distribution:

$$87.5 + 37 \cdot \mathrm{BETA}(0.618, 0.596). \quad (1)$$

This initial distribution was derived by analyzing the observed surgical durations, which ranged between 87.5 and 124.5 minutes, as determined by the Beta distribution parameters. To further refine the data generation process, an intermediate step involved fitting a Beta distribution with parameters Alpha1 = 0.618, Alpha2 = 0.596, Min = 88, and Max = 124, as shown in Fig. 1. This step helped confirm the suitability of the Beta distribution for modeling the variability in surgical durations. Based on this, 100 new samples were simulated, resulting in a refined distribution:

$$88 + 36 \cdot \mathrm{BETA}(0.58, 0.656). \quad (2)$$

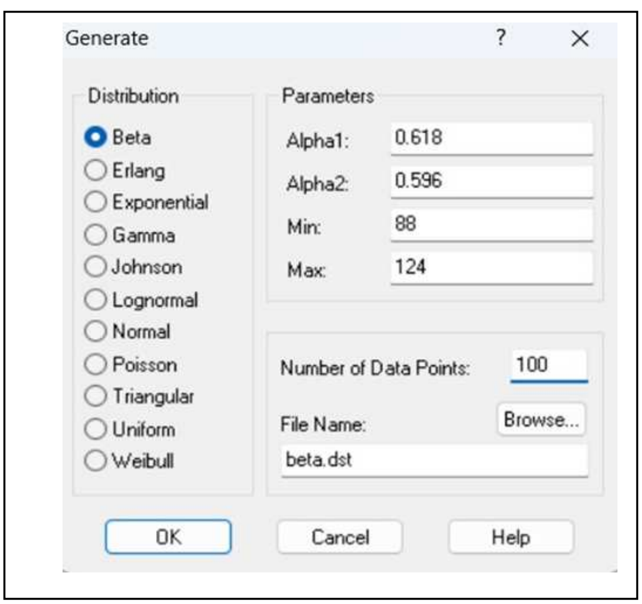

Figure 1. Beta Distribution Fitting in ARENA Input Analyzer

The slight adjustment in parameters reflects an iterative process to better capture the underlying distribution of the data, ensuring that the synthetic samples closely resembled the real-world observations. The final dataset included key variables such as patient age, weight, surgeon grade, assistant grade, and the actual duration of the procedure.

B. Key Variables and Feature Selection

The template is used to format your paper and style the text. All margins, column widths, line spaces, and text fonts are prescribed; please do not alter them. You may note peculiarities. For example, the head margin in this template measures proportionately more than is customary. This measurement and others are deliberate, using specifications that anticipate your paper as one part of the entire proceedings, and not as an independent document. Please do not revise any of the current designations.

To improve the accuracy of predictive models for surgical procedure duration, it is essential to identify the most relevant variables in the dataset. In this study, interviews were conducted with experienced surgeons, anesthesiologists, and other healthcare professionals to determine which factors are likely to influence surgery time. Among the variables discussed, the type of surgical procedure and the mode of anesthesia were highlighted as significant. Since the focus is specifically on cholecystectomy, which is always performed under general anesthesia with a standard theoretical duration of 90 minutes, only applicable features were retained.

Based on insights from operating room staff, several key variables were selected for their potential impact and availability in patient records. The details of these variables are listed in Table I below. These include patient age, the grade of the lead surgeon, the grade of the assistant surgeon, patient weight, the actual duration of the procedure, and its theoretical duration. Surgeon grades were classified as senior (noted as 1) and resident (noted as 2). These variables form the foundation for building accurate predictive models within the surgical department of Mahmoud El Matri Hospital.

C. Data Preprocessing

Data were normalized to ensure scale consistency, outliers were removed, and missing values were addressed via imputation or exclusion, preparing the dataset for machine learning.

TABLE I. VARIABLES AFFECTING THE DURATION OF THE SURGICAL PROCEDURE

Variable	Type
Patient age	Numeric
Surgeon’s grade	Numeric
Assistant surgeon’s grade	Numeric
Patient’s weight	Numeric
Actual duration of the surgical procedure	Numeric
Theoretical duration of the surgical procedure	Numeric

D. Model Development

To enhance the accuracy of surgical duration predictions, particularly for cholecystectomy procedures, this study introduces a hybrid machine learning model that integrates supervised and unsupervised learning techniques. Specifically, it combines Random Forest regression with K-Means clustering. This two-phase model leverages the complementary strengths of both algorithms to improve predictive performance and robustness.

III. PREDICTION MODEL

A. Random Forest

Random Forest (RF), developed by Breiman and Cutler, is an ensemble learning algorithm that combines the predictions of multiple decision trees to deliver accurate and robust results [11]. This method has been widely applied across various domains, including medicine, where it has proven effective for predicting surgical durations due to its ability to handle complex and heterogeneous data [9], [12]. In this study, RF is utilized to predict the duration of cholecystectomy procedures using features such as patient age, weight, and the experience level of the surgical team.

In the context of surgical duration prediction, RF excels at managing the complexity and variability inherent in surgical procedures, which are influenced by numerous factors including patient condition, surgical technique, and the medical team's expertise. The algorithm's ability to handle high-dimensional, noisy datasets while avoiding overfitting makes it particularly well-suited for this task. Additionally, RF provides an internal estimate of prediction accuracy through out-of-bag (OOB) samples, enhancing its generalizability and reliability in clinical settings [2], [7].

The Random Forest algorithm operates through the following steps:

- Randomly select a subset of features and a bootstrap sample of the dataset.
- Construct a decision tree based on the selected subset of features and sample.
- Repeat steps 1 and 2 to create a large number of decision trees, each trained on a different subset of the data and features.
- Aggregate the predictions from all decision trees to produce a final prediction, typically by averaging the outputs in regression tasks.

This process ensures diversity among the decision trees by introducing randomness in both the data sampling (via bootstrapping) and feature selection at each split. By aggregating the results of multiple trees, RF reduces the risk of overfitting and improves the stability and accuracy of predictions, making it a robust choice for modeling the non-linear relationships between surgical features and procedure duration.

Mathematically, the Random Forest algorithm can be described as follows. Let X be the dataset with n instances and m features, and Y be the target variable (surgical duration in this case). The algorithm constructs k decision trees, denoted as $T_1, T_2, \dots, T_k$, each built using a bootstrap sample of the data and a randomly selected subset of features. For a given instance X, the prediction of the i-th tree is $T_i(X)$, representing the predicted value of the target variable Y. The final prediction Y is obtained by aggregating the predictions of all trees, which, for regression, is expressed as the average:

$$\hat{Y} = \frac{1}{k}\sum_{i=1}^{k} T_i(X) \tag{3}$$

One of the key advantages of Random Forest in this application is its ability to provide feature importance scores, which offer insights into the factors most influential in predicting surgical duration. Analysis of the feature importance revealed that surgeon grade and patient weight were among the top contributors, aligning with clinical intuition that more experienced surgeons and heavier patients often require different operative times due to variations in procedural complexity. This interpretability enhances the model's practical utility, as it allows healthcare providers to understand the driving factors behind the predictions and potentially adjust OR planning accordingly.

In the context of surgical duration prediction, Random Forest proves to be an effective and reliable tool. Its capacity to uncover complex, non-linear relationships between surgical features and procedure duration enables precise predictions, even in the presence of heterogeneous and high-dimensional data. Furthermore, RF's robustness makes it suitable for real-time applications, where it can adapt to the dynamic nature of OR scheduling by providing timely and accurate duration estimates. By improving the precision of surgical duration predictions, Random Forest contributes to enhanced OR efficiency, reduced delays, and better resource allocation, ultimately supporting safer and more effective surgical procedures.

B. K-Means Clustering

K-Means is an unsupervised machine learning algorithm that groups data points into a predefined number of clusters based on similarity, making it an effective method for addressing unobserved heterogeneity in surgical data. In this study, K-Means clustering is applied to demographic and clinical features such as patient age, weight, and medical history to identify patterns and group patients with similar profiles. This step helps reduce data heterogeneity by creating more homogeneous subgroups, which can reveal underlying trends that influence surgical durations and improve the subsequent predictive performance of the Random Forest regression model [13]-[14].

The K-Means algorithm operates iteratively to partition the dataset into k clusters by minimizing the variance within each cluster. The process begins by randomly selecting k initial centroids, where k is the predefined number of clusters. Each data point is then assigned to the nearest centroid based on the Euclidean distance, which serves as the similarity metric. After all points are assigned, the algorithm recalculates the centroid of each cluster as the mean of all points within that cluster. This process continues, with points being reassigned and centroids updated, until it reaches convergence, when the assignments remain unchanged or the change in the objective function becomes insignificant.

The primary objective of K-Means is to minimize the within-cluster sum of squared errors, also known as the inertia. Mathematically, given a dataset $X=\{x_1,x_2,\ldots,x_n\}$ of n data points and a set of k centroids $C=\{c_1,c_2,\ldots,c_k\}$, the optimization goal is to find the cluster assignments and centroids that minimize the following objective function:

$$arg\min_{C}\sum_{i=1}^{n}\min_{c_j\in C}\left\|x_i-c_j\right\|^2 \tag{4}$$

Where x_i represents the i-th data point, c_j represents the j-th centroid, $\|x_i-c_j\|^2$ denotes the squared Euclidean distance between the data point x_i and the centroid c_j.

This equation ensures that each data point is assigned to the cluster whose centroid is closest, thereby minimizing the total variance within clusters. The iterative nature of K-Means ensures that the algorithm converges to a local minimum, although the quality of the solution can depend on the initial choice of centroids. To mitigate this, multiple runs with different initializations were performed, and the solution with the lowest inertia was selected.

By applying K-Means clustering, the heterogeneity in the surgical data was significantly reduced, as patients within each cluster shared similar demographic and clinical characteristics. This homogeneity allowed the subsequent Random Forest regression model to focus on learning patterns within more cohesive groups, ultimately enhancing the overall predictive accuracy of the hybrid model. The cluster labels assigned by K-Means were added as an additional feature to the dataset, providing the Random Forest model with contextual information about patient similarity that improved its ability to generalize across diverse cases.

C. Hybrid Model Architecture

The hybrid model integrates K-Means clustering and Random Forest regression in a two-step approach:

- Clustering Step: Patients were grouped based on age, weight, and medical history using Euclidean distance. K-Means minimizes within-cluster variance, iteratively updating centroids until convergence. This step reduces data heterogeneity by identifying similar patient profiles.
- Random Forest Regression: Applied post-clustering, RF constructs multiple decision trees on bootstrapped subsets of features and data, averaging predictions to mitigate overfitting. It uses selected features to estimate duration.

By first reducing data heterogeneity and then applying a robust predictive algorithm, the hybrid model improves prediction accuracy and generalizability compared to standalone models.

D. Implementation Strategy

The dataset was split into training (80%) and testing (20%) sets. K-Means clustering was applied to the training data to assign cluster labels, which were then added as input features for the RF regression model. Model performance was evaluated using key metrics including Mean Absolute Error (MAE), Root Mean Squared Error (RMSE), and R-squared (R^2).

IV. Results and Discussions

By combining unsupervised clustering to reduce data heterogeneity with supervised regression to capture complex patterns, the hybrid model aims to provide precise and generalizable predictions. The following subsections detail the model's performance through quantitative metrics, visual comparisons, and benchmarking against traditional and standalone approaches, highlighting its potential as a practical tool for OR scheduling.

A. Quantitative Model Comparison

To rigorously evaluate the predictive performance of the hybrid model, a quantitative comparison was conducted against two standalone machine learning algorithms: Random Forest (RF) and K-Means. These models were selected as baselines because RF represents a robust supervised learning approach commonly used for regression tasks, while K-Means, though primarily a clustering algorithm, was adapted here as a simplistic predictor by assigning the mean duration of each cluster as the predicted value for its members. The comparison focuses on regression metrics, Root Mean Square Error (RMSE), Mean Absolute Error (MAE), and the coefficient of determination (R^2), to assess prediction accuracy, as well as classification metrics (precision, recall, and F1 score) for categorizing surgeries into short (< 90 minutes) and long (> 120 minutes) durations, which are clinically relevant thresholds for OR scheduling.

The results of this comparison are summarized in Table II, with all models evaluated on the same test set. The hybrid model achieves an RMSE of 0.4509 minutes, slightly outperforming the standalone Random Forest model (0.4566 minutes) and significantly surpassing K-Means (100.18 minutes). Similarly, the hybrid model's MAE of 0.36 minutes is lower than that of Random Forest (0.37 minutes) and K-Means (99.26 minutes), indicating a higher average prediction accuracy. The R^2 score of the hybrid model, at 0.9987, demonstrates that it explains 99.87% of the variance in surgical duration, compared to 0.997 for Random Forest and a negative R^2 of -58.99 for K-Means, highlighting K-Means' unsuitability as a standalone predictor due to its inability to model continuous outcomes effectively.

The superior performance of the hybrid model in regression metrics can be attributed to its two-step architecture. The K-Means clustering step reduces data heterogeneity by grouping similar patients, allowing the Random Forest component to focus on learning patterns within more homogeneous subgroups. This synergy results in more precise predictions, as evidenced by the lower RMSE and MAE compared to the standalone Random Forest, which lacks the clustering pre-step and thus struggles with the full dataset's variability. The K-Means model's poor performance underscores its limitation as a regression tool, as it oversimplifies the prediction task by assigning cluster means without accounting for feature-specific variations within clusters.In addition to regression metrics, the hybrid model was evaluated for its ability to classify surgeries into short (< 90 minutes) and long (> 120 minutes) durations, which are critical thresholds for OR scheduling and resource allocation. Surgeries shorter than 90 minutes typically require less intensive preparation and recovery, while those longer than 120 minutes may necessitate additional staffing or equipment. The hybrid model achieves perfect classification performance, with precision, recall, and F1 scores of 1.0 for both short and long surgeries. This indicates that the model correctly identifies all instances of short and long surgeries without false positives or false negatives. In contrast, the standalone Random Forest and K-Means models were not evaluated for classification, as their primary design (regression for RF, clustering for K-Means) does not naturally lend itself to this task without additional modifications, such as thresholding their continuous predictions.

The scatter plot in Fig. 2 complements these quantitative findings by visually comparing the predicted versus actual surgical durations for the hybrid model. The tight clustering of points along the ideal line (where predicted equals actual) reinforces the model's high accuracy, with minimal deviations even for edge cases. This visual evidence, combined with the regression and classification metrics, underscores the hybrid model's robustness across different evaluation dimensions.

Overall, the quantitative comparison demonstrates the hybrid model's superior predictive capability compared to standalone Random Forest and K-Means models. Its ability to achieve lower errors, explain nearly all variance in the data, and perfectly classify short and long surgeries highlights its potential as a reliable tool for OR scheduling. The statistically significant improvement over Random Forest further validates the value of integrating clustering with regression, offering a practical and effective solution for surgical duration prediction.

B. Benchmarking Against Traditional Models

To further validate our results, we compared the hybrid model with more traditional approaches: Decision Tree and Linear Regression. As shown in Table III, the hybrid model consistently achieves the lowest RMSE of 0.45 minutes, confirming its superior performance. Compared to both the decision tree and linear regression models, the hybrid model provides significantly more accurate predictions. Overall, these results underscore the hybrid model's high reliability, precision, and general effectiveness in predicting surgical durations across different scenarios.

C. Generalization to New Data

To assess how well the hybrid model performs on unseen data, we followed a structured evaluation process:

1. Insert new input data for which the models had not previously been trained.
2. Use these inputs to predict surgical durations using the hybrid model and other baseline models.
3. Compare the prediction errors using MAE, RMSE, and R^2.
4. Calculate average performance metrics for each model to evaluate overall accuracy.

The analysis of the overall prediction performance in Table IV reveals that the hybrid model offers the highest accuracy, with a mean absolute error (MAE) of just 12.59 minutes. This superior performance suggests that the hybrid model is the most effective at accurately predicting surgical procedure durations. In comparison, other models such as Random Forest, Decision Tree, KMeans, and Linear Regression exhibit higher MAE values, indicating lower precision in their predictions. Based on these results, it is reasonable to conclude that the hybrid model is the best choice for predicting surgical procedure durations.

TABLE II. PERFORMANCE COMPARISON WITH RANDOM FOREST AND KMEANS

Metric	Hybrid Model	Random Forest	KMeans
RMSE	0.4509 minutes	0.4566 minutes	100.18 minutes
MAE	0.36 minutes	0.37 minutes	99.26 minutes
R^2 Score	0.9987	0.997	-58.99
Precision (< 90 min)	1	N/A	N/A
Recall (< 90 min)	1	N/A	N/A
F1 Score (< 90 min)	1	N/A	N/A
Precision (> 120 min)	1	N/A	N/A
Recall (> 120 min)	1	N/A	N/A
F1 Score (> 120 min)	1	N/A	N/A

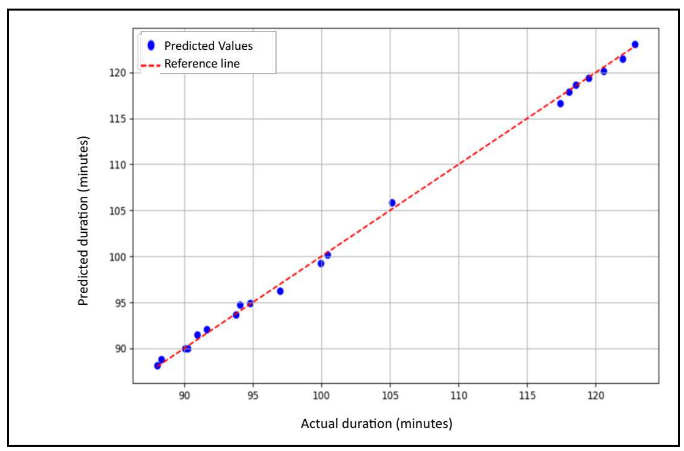

Figure 2. Comparison of Predicted vs. Actual Values

TABLE III. PERFORMANCE COMPARISON WITH DECISION TREE AND LINEAR REGRESSION MODELS

Metric	Hybrid Model	Decision Tree	Linear Regression
RMSE	0.45 minutes	0.57 minutes	1.96 minutes
MAE	0.37 minutes	0.44 minutes	1.44 minutes
R^2 Score	0.9987	0.9980	0.97

TABLE IV. OVERALL PREDICTION PERFORMANCE

Model	MAE (minutes)
Hybrid Model	12.59
Random Forest	14.18
KMeans	104.75
Decision Tree	14.14
Linear Regression	15.07

V. Conclusion

This study presents a hybrid machine learning model that integrates Random Forest regression and K-Means clustering to predict surgical procedure durations, with a focus on cholecystectomy cases at Mahmoud El Matri Hospital in Tunis, Tunisia. The model leverages real-world data enriched with synthetic samples, incorporating key variables such as patient age, weight, and surgeon experience to deliver highly accurate predictions. Evaluation results highlight its exceptional performance. Compared to standalone models like Random Forest, K-Means, decision trees, and linear regression, the hybrid approach consistently outperforms, demonstrating its ability to capture complex patterns and reduce prediction errors effectively.

The success of this model underscores the value of combining supervised and unsupervised learning techniques to address the challenges of surgical duration prediction. By segmenting patients into meaningful clusters and applying robust regression, the hybrid model not only enhances accuracy but also offers a practical solution for optimizing operating room scheduling. This advancement has the potential to minimize delays, improve resource allocation, and enhance patient care, aligning with the broader goals of healthcare efficiency.

While these findings are promising, limitations remain, including reliance on synthetic data and the need for validation across diverse surgical procedures and settings. Future research should focus on testing the model with larger, real-time datasets, extending its application to other types of surgeries, and exploring additional features to further refine its predictive power. Ultimately, this hybrid model represents a significant step toward data-driven operating room management, offering a scalable and adaptable tool to support healthcare providers in delivering timely and cost-effective care.

References

[1] M. Eshghali, D. Kannan, N. Salmanzadeh-Meydani, and A. M. Esmaieeli Sikaroudi, “Machine learning based integrated scheduling and rescheduling for elective and emergency patients in the operating theatre,” Ann. Oper. Res., vol. 332, no. 1, pp. 989–1012, 2024.

[2] L. Luo, C. Liu, L. Feng, S. Zhao, and R. Gong, “A random forest and simulation approach for scheduling operation rooms: Elective surgery cancelation in a Chinese hospital urology department,” Int. J. Health Planning Manage., vol. 33, no. 4, pp. 941–966, 2018.

[3] A. Erekat, G. Servis, S. C. Madathil, and M. T. Khasawneh, “Efficient operating room planning using an ensemble learning approach to predict surgery cancellations,” IISE Trans. Healthcare Syst. Eng., vol. 10, no. 1, pp. 18–32, 2020.

[4] T. G. Smith et al., “Creating a practical transformational change management model for novel artificial intelligence–enabled technology implementation in the operating room,” Mayo Clin. Proc. Innov. Qual. Outcomes, vol. 6, no. 6, pp. 584–596, 2022.

[5] V. Bellini et al., “Artificial intelligence in operating room management,” J. Med. Syst., vol. 48, no. 1, p. 19, 2024.

[6] M. Fairley et al., “Enhancing OR efficiency with optimization and ML,” Health Care Manag. Sci., vol. 22, no. 4, pp. 756–767, 2019.

[7] K. Lin, Y. Hu, and G. Kong, “Predicting in-hospital mortality of patients with acute kidney injury in the ICU using random forest model,” Int. J. Med. Informatics, vol. 125, pp. 55–61, 2019.

[8] L. Luo, H. Liu, X. Hou, and Y. Shi, “Machine learning methods for surgery cancellation,” in Proc. 13th Int. Conf. Service Syst. Service Manage. (ICSSSM), Kunming, China, Jun. 2016, pp. 1–4.

[9] Z. ShahabiKargar, S. Khanna, N. Good, A. Sattar, J. Lind, and J. O’Dwyer, “Predicting procedure duration to improve scheduling of elective surgery,” in PRICAI 2014: Trends in Artificial Intelligence, Gold Coast, QLD, Australia, Dec. 2014, pp. 998–1009.

[10] S. Amrollahibiouki and Y. Beauregard, "Surgical duration prediction by developing a hybrid model (Using machine learning techniques)," in CIGI Qualita MOSIM 2023, Trois-Rivières, Québec, Canada, 2023.

[11] L. Breiman, “Random forests,” Mach. Learn., vol. 45, no. 1, pp. 5–32, 2001.

[12] V. Riahi et al., “Improving preoperative prediction of surgery duration,” BMC Health Services Res., vol. 23, no. 1, p. 1343, 2023

[13] G. Spini, M. van Heesch, T. Veugen, and S. Chatterjea, “Private hospital workflow optimization via secure k-means clustering,” J. Med. Syst., vol. 44, no. 1, p. 8, 2020.

[14] L. W. Santoso, A. Sudiarso, N. A. Masruroh, and M. K. Herliansyah, “Cluster analysis to determine the priority of operating room scheduling,” in AIP Conf. Proc., vol. 1977, no. 1, Jun. 2018.

[15] J. B. Park, G. H. Roh, K. Kim, and H. S. Kim, “Development of predictive model of surgical case durations using machine learning approach,” J. Med. Syst., vol. 49, no. 1, pp. 1–11, 2025.

[16] L. Luo, C. Liu, L. Feng, S. Zhao, and R. Gong, “A random forest and simulation approach for scheduling operation rooms: Elective surgery cancelation in a Chinese hospital urology department,” Int. J. Health Planning Manage., vol. 33, no. 4, pp. 941–966, 2018.